

Coupling the inflaton to an expanding aether

William Donnelly* and Ted Jacobson†

*Center for Fundamental Physics, Department of Physics
University of Maryland, College Park, MD 20742-4111*

We consider a Lorentz-violating theory of inflation consisting of Einstein-aether theory with a scalar inflaton coupled bilinearly to the expansion of the aether. We determine the conditions for linearized stability, positive energy and vanishing of preferred-frame post-Newtonian parameters, and find that all these conditions can be met. In homogeneous and isotropic cosmology, the inflaton-aether expansion coupling leads to a driving force on the inflaton that is proportional to the Hubble parameter. This force affects the slow-roll dynamics, but still allows for a natural end to inflation.

PACS numbers: 04.50.Kd, 11.30.Cp, 98.80.Cq

I. INTRODUCTION

The space-time continuum may be only a low-resolution approximation to a more structured, discrete plenum, with a UV cutoff on degrees of freedom. This structure might be revealed in high energy collisions or subtle effects of long distance propagation, or discrete symmetry violations. But it might also manifest itself at a macroscopic level as a result of the expansion of the Universe. The expansion would presumably require that new “atoms of space-time” together with new degrees of freedom are created as time goes on. The aim of this paper is to explore possible consequences of this scenario using an effective field theory description.

A UV cutoff is inconsistent with locality and Lorentz symmetry, so at least one of these properties would have to fail if such a cutoff exists. In this paper we study a model theory in which locality is preserved, but Lorentz symmetry fails. To accommodate this scenario in an effective field theory that preserves general covariance and the successes of general relativity, one needs to incorporate a Lorentz-violating dynamical field into the theory. Einstein-aether theory provides a straightforward approach to doing this. In that theory, the local structure of space-time is described by a metric tensor, as in general relativity, but also by a dynamical unit timelike vector field u^a , called the aether. The aether defines a preferred rest frame at each point of space-time, but preserves rotational symmetry in that frame. This structure would suffice to accommodate a local UV cutoff. For a Lorentz violating theory like this to be viable, there must ultimately be a reason why conspicuous Lorentz violation does not infect the low-energy matter action in flat space-time. For the purposes of this paper, we will simply assume that such a reason exists.

Just as the Einstein-Hilbert action is (besides a cosmological constant term) the unique lowest order covariant term in a derivative expansion of the action, the action

for “pure” Einstein-aether theory consists, in addition, of the four independent two-derivative terms involving the metric and the aether. Direct coupling of the aether to matter would entail local Lorentz violation and is severely constrained by observational bounds [1].

Our primary goal is to use Einstein-aether theory to model possible consequences associated with the growth of the vacuum in cosmology. In particular, we study the leading order coupling of a scalar inflaton field φ to the local expansion rate. Such a term cannot appear in a purely metric theory because the local expansion rate, i.e. the Hubble parameter H , cannot be constructed from the metric in a covariant way. However, given the aether field u^a , the local expansion θ relative to the preferred frame is a space-time scalar,

$$\theta \equiv \nabla_a u^a. \quad (1)$$

(In a homogeneous and isotropic cosmology, θ is related to the Hubble parameter by $\theta = 3H$.) Any scalar operator might couple to θ , but here we shall consider just a (still mysterious) scalar inflaton, since that would presumably dominate during inflation. This of course leaves open the possibility that scalar operators composed of matter fields could couple to θ . At late times θ is everywhere small relative to particle physics scales, so θ couplings would not lead to any observable Lorentz violation in particle physics.

We shall consider a generic coupling of the inflaton to the expansion, but the lowest dimension (and therefore presumably dominant) coupling takes the form

$$\theta\varphi = -u^a \nabla_a \varphi + \text{total derivative}. \quad (2)$$

This peculiar leading order term has not, as far as we know, been considered previously. It apparently entails a violation of time-reversal symmetry, although if the rest of the inflaton action is even in φ then time-reversal symmetry is preserved when accompanied with $\varphi \rightarrow -\varphi$. We first examine the impact of this new coupling on all of the noncosmological theoretical and observational constraints on the theory, and then we study the modified inflationary dynamics. A key consequence of this term is that the expansion θ acts as an external force, which can either slow down or accelerate the evolution of the

*Electronic address: wdonnell@umd.edu

†Electronic address: jacobson@umd.edu

inflaton. Its effect on the primordial fluctuations has not yet been fully investigated.

Although the particular coupling we examine may be new, much work has already been done on cosmology in Einstein-aether theory. Inflation with a preferred time-like vector was first considered by Gasperini [2]. More recently, authors have computed the effect of the aether on the spectrum of primordial perturbations [3–6], with no direct coupling between the aether and the inflaton. A certain kind of aether-inflaton coupling has been considered in which the kinetic terms in the aether action are multiplied by functions of the scalar field [7, 8]. In this model the effective gravitational constant depends on the inflaton, leading to the possibility of a repulsive phase that significantly changes the dynamics of inflation. A similar model with a nonnormalized vector field playing the role of the aether and scalar fields has been proposed as a solution of the cosmological constant problem by a similar mechanism [9]. Cosmology with a Lagrangian that is an arbitrary function of the usual Einstein-aether Lagrangian has also been considered as a possible source of dark energy [10, 11].

This paper is organized as follows. In Sec. II the theory is defined and the equations of motion are given. In Secs. III, IV, and V we work out the properties of linearized spin-0 perturbations and the conditions for their stability, absence of vacuum Cherenkov radiation, and energy positivity. This is necessary in order to establish the conditions for viability of the model with aether-scalar coupling. In Sec. VI these conditions are combined with those for the spin-1 and spin-2 perturbations of pure Einstein-aether theory, the effect of the new coupling on the parametrized post-Newtonian (PPN) expansion is deduced, and some additional observational constraints are discussed. In Sec. VII the homogeneous, isotropic cosmological dynamics is investigated, and we conclude with a summary and some further remarks in Sec. VIII.

II. ACTION AND EQUATIONS OF MOTION

Einstein-aether theory is a theory of a metric g_{ab} of signature $(+,-,-,-)$ and a unit timelike vector field u^a [12]. This theory is defined by the Lagrangian

$$L_a = \frac{-M^2}{2} (R + K^{ab}_{cd} \nabla_a u^c \nabla_b u^d + \lambda(u^a u_a - 1)), \quad (3)$$

where $M \equiv (8\pi G)^{-\frac{1}{2}}$ is the reduced Planck mass. In general relativity, the sign of G can be fixed by requiring that gravitational waves carry positive energy density. The same turns out to be true in Einstein-aether theory, hence we will assume from the outset that $G > 0$, or equivalently $M^2 > 0$. The three terms in the action are the Ricci scalar for g_{ab} , the Einstein-aether kinetic term, and a Lagrange multiplier term that forces u^a to be a timelike unit vector. The tensor K^{ab}_{cd} is given by

$$K^{ab}_{cd} = c_1 g^{ab} g_{cd} + c_2 \delta_c^a \delta_d^b + c_3 \delta_d^a \delta_c^b + c_4 u^a u^b g_{cd}, \quad (4)$$

where $c_{1,2,3,4}$ are the dimensionless free parameters of Einstein-aether theory.

The theory defined by the Lagrangian (3) has a Newtonian limit with a modified Newton's constant

$$G_N = \kappa_N G, \quad \kappa_N = \frac{2}{2 - c_{14}}. \quad (5)$$

The post-Newtonian expansion will be discussed in Sec. VI B. The Friedmann equation for spatially flat homogeneous cosmologies is also the same as in general relativity, but with a renormalized Newton's constant G_C in place of G [13, 14],

$$G_C = \kappa_C G, \quad \kappa_C = \frac{2}{2 + c_{13} + 3c_2}. \quad (6)$$

The spatial curvature term in the Friedmann equation is also renormalized.

We introduce to the theory a scalar inflaton φ with a potential that depends both on φ and the expansion of the aether,

$$L_\varphi = \frac{1}{2} \nabla_a \varphi \nabla^a \varphi - V(\theta, \varphi). \quad (7)$$

For special cases of the potential $V(\theta, \varphi)$ the cosmological dynamics of our model overlaps with other models of Lorentz-violating inflation. To see this, we note that in a Friedmann-Lemaître-Robertson-Walker metric with the aether aligned to the cosmological rest frame,

$$\nabla_a u_b = H(g_{ab} - u_a u_b), \quad (8)$$

where $H = \dot{a}/a$ is the Hubble parameter. It follows that the aether kinetic term in (3) is proportional to θ^2 . In the model of Kanno and Soda [7], $c_{1,2,3,4}$ are allowed to depend on a scalar field φ . This will have the same cosmological dynamics as our model with a potential of the form $V(\theta, \varphi) = f(\varphi)\theta^2$. The model of Zlosnik et. al. [10] does not have a scalar field, but its Lagrangian contains an arbitrary function of the aether kinetic term. This cosmology corresponds to a potential of the form $V(\theta, \varphi) = f(\theta^2)$. This equivalence holds only for the background dynamics; the linearized perturbations of these models will generally differ.

The equation of motion for φ is

$$\square \varphi + V_\varphi = 0, \quad (9)$$

where V_φ denotes the partial derivative $\partial V / \partial \varphi$. The equation of motion for u is

$$\nabla_b K^b_a = \lambda u_a + c_4 (u^c \nabla_c u_b) \nabla_a u^b - M^{-2} \nabla_a V_\theta, \quad (10)$$

where $K^b_a \equiv K^{ab}_{cd} \nabla_a u^c$. The component of Eq. (10) parallel to u^a determines λ . Finally we have the Einstein equation:

$$G_{ab} = S_{ab} + M^{-2} T_{ab}, \quad (11)$$

where S_{ab} is the energy-momentum tensor for the aether in the absence of the scalar field [15]

$$\begin{aligned} S_{ab} = & \nabla_c \left(K_{(a}^c u_{b)} - K_{(a}^c u_{b)} - K_{(ab)} u^c \right) \\ & + c_1 (\nabla_c u_a \nabla^c u_b - \nabla_a u_c \nabla_b u^c) \\ & + c_4 (u^c \nabla_c u_a) (u^d \nabla_d u_b) \\ & + (u^c \nabla_d K_c^d - c_4 (u^c \nabla_c u^e) (u^d \nabla_d u_e)) u_a u_b \\ & + \frac{1}{2} g_{ab} (K_d^c \nabla_c u^d), \end{aligned} \quad (12)$$

round brackets denoting symmetrization. The matter energy momentum tensor T_{ab} is obtained by varying the matter action with respect to the metric. Accounting for the contribution of the Lagrange multiplier term the energy-momentum tensor can be written in terms of the Lagrangian as

$$T_{ab} = 2 \frac{\delta L}{\delta g^{ab}} + u^c \frac{\delta L}{\delta u^c} u_a u_b - L g_{ab}. \quad (13)$$

For the scalar Lagrangian (7) this formula gives

$$\begin{aligned} T_{ab} = & \nabla_a \varphi \nabla_b \varphi - \left(\frac{1}{2} \nabla_a \varphi \nabla^a \varphi - V + \theta V_\theta \right) g_{ab} \\ & + (u^c \nabla_c V_\theta) (u_a u_b - g_{ab}). \end{aligned} \quad (14)$$

Note that terms in the action linear in θ do not contribute to $V - \theta V_\theta$, hence they contribute only to the part of the stress tensor proportional to the spatial metric $u_a u_b - g_{ab}$. This can be traced to the contributions from the Lagrange multiplier term in the action, but it can also be understood more directly from the fact that the metric enters such terms via the combination $\sqrt{-g}\theta = \partial_a(\sqrt{-g}u^a)$. In order to maintain the unit norm condition $g_{ab}u^a u^b = 1$, a metric variation must be accompanied by an aether variation $\delta_{\parallel} u^a = -\frac{1}{2}(\delta g_{mn} u^m u^n)u^a$. The combined δg^{ab} and $\delta_{\parallel} u^a$ variations yield

$$\delta(\sqrt{-g}u^a) = \frac{1}{2}\sqrt{-g}u^a(-g_{mn} + u_m u_n)\delta g^{mn}. \quad (15)$$

The corresponding contribution to the stress tensor is thus proportional to the spatial metric, and corresponds to an isotropic pressure in the aether frame.

III. LINEARIZED PERTURBATIONS AND DISPERSION RELATIONS

We consider linearized perturbations about the flat space-time solution with a constant aether field

$$g_{ab} = \eta_{ab}, \quad u^a = \delta_0^a, \quad \varphi = 0. \quad (16)$$

To ensure that (16) is a solution of the equations of motion, the potential must satisfy

$$V(0,0) = V_\varphi(0,0) = 0. \quad (17)$$

In order to study linear perturbations, the potential is expanded to second order in the fields

$$V(\theta, \varphi) \approx \frac{1}{2}m^2\varphi^2 + \mu M\theta\varphi, \quad (18)$$

where m is the mass of the scalar field. The constant μ has dimensions of mass, and we include a factor of M to compensate for the canonical normalization of the aether perturbations. The expansion (18) does not include a term linear in θ , because such a term is a total divergence and therefore does not affect the classical equations of motion. The term proportional to θ^2 is also omitted, because it can be absorbed into the c_2 term in the Einstein-aether Lagrangian.

Because the background solution (16) is invariant under spatial rotations, we can decompose the perturbations into irreducible representations of SO(3). The propagating degrees of freedom consist of a spin-2 graviton, a spin-1 aether-metric wave, a spin-0 aether-metric wave, and a spin-0 inflaton perturbation. At the linearized level, the presence of the scalar field affects only the spin-0 modes. This is because the term $\theta\varphi$ in the action depends only on the timelike and longitudinal parts of the aether field, both of which have spin-0.

We will adopt the variables and gauge choice of Foster [16]. In this parametrization, the seven spin-0 degrees of freedom are labeled $h_{00}, \gamma, \phi, f, w^0, \nu, \varphi$, and a general spin-0 perturbation of the metric and aether can be written as

$$g_{ab} = \begin{bmatrix} 1 + h_{00} & \partial_i \gamma \\ \partial_i \gamma & -1 + \partial_i \partial_j \phi + \frac{1}{2}(\Delta f \delta_{ij} - \partial_i \partial_j f) \end{bmatrix}, \quad (19)$$

$$u^a = [1 + w^0, \partial_i \nu]. \quad (20)$$

Here 0 labels the time coordinate, latin indices $i, j, \dots = 1, 2, 3$ are spatial, and $\Delta = \delta^{ij} \partial_i \partial_j$. We can simplify the equations of motion by using the gauge freedom to set $\nu = \gamma = 0$ [16]. Additionally, we can eliminate w^0 using the constraint $u^a u_a = 1$, giving

$$w^0 = -\frac{1}{2}h_{00}. \quad (21)$$

With these choices, and taking perturbations to have the space-time dependence $e^{i(kx - \omega t)}$ in terms of the flat space coordinates (t, x, y, z) , the metric and aether take the form

$$g_{ab} = \text{diag}(1 + h_{00}, -1 - k^2\phi, -1 - \frac{1}{2}k^2f, -1 - \frac{1}{2}k^2f), \quad (22)$$

$$u^a = (1 - \frac{1}{2}h_{00}, 0, 0, 0). \quad (23)$$

In terms of these fields the linearized versions of the scalar field equation, the (0,0) Einstein equation, the (1,1) Einstein equation and the sum of the (2,2) and (3,3) Einstein equations are

$$0 = -\omega^2\varphi + k^2\varphi + m^2\varphi - \frac{1}{2}k^2i\omega\mu M(f + \phi), \quad (24)$$

$$0 = -k^2f - c_{14}h_{00}, \quad (25)$$

$$0 = (1 + c_2)k^2\omega^2f + c_{123}k^2\omega^2\phi + 2\mu i\omega\varphi/M, \quad (26)$$

$$\begin{aligned} 0 = & (1 + c_{13} + 2c_2)k^2\omega^2f \\ & - k^2(k^2f + 2h_{00} - 2(1 + c_2)\omega^2\phi) + 4\mu i\omega\varphi/M. \end{aligned} \quad (27)$$

where we have adopted the notation $c_{14} = c_1 + c_4$, $c_{123} = c_1 + c_2 + c_3$, etc.

In the absence of the scalar field the linearized modes all satisfy a linear dispersion relation of the form $\omega = s_i k$ where s_i , $i = 0, 1, 2$ is a speed depending only on the spin of the mode. This is because the aether terms in the Lagrangian all contain two derivatives. The scalar field action contains terms with fewer than two derivatives, so in general the spin-0 modes are dispersive. The dispersion relation is obtained by taking the determinant of Eqs. (24)-(27) considered as a matrix equation. This can be simplified to

$$(\omega^2 - k^2 - m^2)(\omega^2 - s_0^2 k^2) + \frac{3\kappa_C}{2}\mu^2\omega^2 - \frac{s_0^2}{c_{123}}\mu^2 k^2 = 0, \quad (28)$$

where κ_C is given by (6) and s_0 is the wave speed of the spin-0 mode when $\mu = 0$ [17]

$$s_0^2 = \frac{(2 - c_{14})c_{123}}{(1 - c_{13})(2 + c_{13} + 3c_2)c_{14}}. \quad (29)$$

The fact that there are just two, rather than four, propagating spin-0 modes is a result of diffeomorphism invariance, and is reflected by the fact that Eq. (28) is quadratic in ω^2 and therefore admits two solutions. In general these modes are superpositions of the spin-0 parts of the metric, aether and scalar fields. Their polarizations can be obtained by solving Eqs. (24)-(27) with fixed ω and k ; they depend on the wavenumber and are valid only in our chosen gauge.

In the homogeneous limit $k \rightarrow 0$ the dispersion relation has the form

$$\omega^2(\omega^2 - m^2 + \frac{3}{2}\kappa_C\mu^2) = 0, \quad (30)$$

which shows that one mode remains gapless, while the mass of the other mode is modified by the μ term.

IV. LINEARIZED STABILITY AND CHERENKOV CONSTRAINT

In what follows we will consider stability of the classical theory linearized about flat space-time. Although this is not sufficient to guarantee stability of the full interacting quantum theory, or even of the linearized theory about a curved background, it is a natural physical condition to impose and will lead to constraints on the parameters $c_{1,2,3,4}, m, M, \mu$. In the following three sections we will show that the linearized theory is stable provided μ is not too large compared to m , and the aether parameters lie within the allowed parameter range for Einstein-aether theory.

To say a theory is stable is to say that regular “initial data” does not grow exponentially in “time”. This notion presumes that the theory admits an initial value formulation with respect to some foliation of the space-time by surfaces, called Cauchy surfaces. If the theory is stable

with respect to one fiducial such foliation, then that suffices to establish stability from a physical point of view. This is because regular initial data on any other Cauchy surface can be evolved back to the fiducial Cauchy surface, where it will again be regular, hence the stability of its complete evolution is assured.

In the case of linearized Einstein-aether theory, the field equations are of hyperbolic type, but not with respect to the space-time metric, since different modes travel at different speeds. Nevertheless there is a maximum speed for the modes, and this determines the class of time functions whose constant-time surfaces are Cauchy surfaces for the full set of equations of motion. In particular, the surfaces orthogonal to the aether vector in a Minkowski background can serve as Cauchy surfaces for the purpose of establishing stability.

We should note that Carroll et. al. [18] have proposed a different criterion for stability, namely that linearized perturbations that are oscillatory on any Lorentz time slice should not grow with respect to the corresponding time coordinate, and they showed that this criterion is strong enough to rule out Einstein-aether theory for most values of the parameters $c_{1,2,3,4}$. We believe that this criterion is not the physically relevant one, because it corresponds to imposing stability with respect to a time function whose constant-time surfaces are not Cauchy surfaces for the equations. This issue will be discussed more fully in a forthcoming publication [19].

We therefore impose stability in the aether frame. This means that the dispersion relation (28) must have two real solutions ω_{\pm} for every real wavenumber k . Equation (28) is a quadratic equation in ω^2 with real coefficients; its roots ω_{\pm}^2 are real provided the discriminant is positive. This is necessary but not sufficient, since ω_{\pm}^2 must be positive in order for ω_{\pm} to be real. We will use the fact that two real numbers ω_{\pm}^2 are both positive if and only if the sum $\omega_{+}^2 + \omega_{-}^2$ and product $\omega_{+}^2\omega_{-}^2$ are both positive. This will be convenient to impose since the sum and product of roots can be read directly from the coefficients of the dispersion relation. In what follows we will work backward, first requiring the sum and product roots to be positive and then returning to require the discriminant to be positive.

The sum and product of the roots are given by

$$\omega_{+}^2 + \omega_{-}^2 = (1 + s_0^2)k^2 + (m^2 - \frac{3}{2}\kappa_C\mu^2), \quad (31)$$

$$\omega_{+}^2\omega_{-}^2 = s_0^2k^4 + (m^2 - \mu^2/c_{123})s_0^2k^2. \quad (32)$$

These will be non-negative for all k if and only if

$$s_0^2 \geq 0, \quad (33)$$

$$\kappa_C\mu^2 \leq 2m^2/3, \quad (34)$$

$$\mu^2/c_{123} \leq m^2. \quad (35)$$

Finally, we impose the condition that the roots ω_{\pm}^2 of the quadratic (28) are both real. Thus we require that

the discriminant be positive,

$$(1 - s_0^2)^2 k^4 + 2[(1 + s_0^2)(m^2 - \frac{3}{2}\kappa_C\mu^2) - 2(m^2 - \mu^2/c_{123})s_0^2]k^2 + (m^2 - \frac{3}{2}\kappa_C\mu^2)^2 \geq 0. \quad (36)$$

The discriminant (36) is of the form $ak^4 + bk^2 + c$ with $a, c \geq 0$. It will be positive if $b \geq 0$ or $b^2 - 4ac \leq 0$, in other words if $b \geq -2\sqrt{ac}$. This condition is

$$(1 + s_0^2 + |1 - s_0^2|)(m^2 - \frac{3}{2}\kappa_C\mu^2) - 2(m^2 - \mu^2/c_{123})s_0^2 \geq 0, \quad (37)$$

where we have made use of (34) to simplify the expression. This inequality can be simplified into two cases depending on the sign of $1 - s_0^2$:

$$1/c_{123} \geq 3\kappa_C/2 \quad \text{if } s_0 \geq 1, \quad (38)$$

$$m^2(1 - s_0^2) \geq (3\kappa_C/2 - s_0^2/c_{123})\mu^2 \quad \text{if } s_0 < 1. \quad (39)$$

We now consider a further physical constraint on the theory: If the phase velocity of scalar aether waves is less than the speed of light, then highly energetic particles can Cherenkov radiate in vacuum. Observations of cosmic rays strongly constrain this behavior [20]. We therefore require the phase velocity of aether waves to be ≥ 1 for all k .

First we show that if a mode has phase velocity ≥ 1 for some k then its phase velocity remains ≥ 1 for all k . If a mode becomes subluminal at a particular wavenumber k it would have to satisfy the dispersion relation with $\omega^2 = k^2$,

$$[m^2(s_0^2 - 1) + (\frac{3}{2}\kappa_C - s_0^2/c_{123})\mu^2]k^2 = 0. \quad (40)$$

If this equation holds for any $k > 0$ then it holds for all k . This means that if a given mode has speed ≥ 1 for any k , it has speed ≥ 1 for all k .

It is therefore sufficient to enforce that the modes are superluminal in the limit $k \rightarrow \infty$. To find the phase velocities in this limit, we express (28) as a polynomial in ω and keep only the dominant power of k in each coefficient. The resulting mode speeds are simply 1 and s_0 , so the Cherenkov constraint is simply $s_0 \geq 1$, as it is in pure Einstein-aether theory.

Thus when the Cherenkov constraint is imposed, (38) is the condition for stability. Equations (34), (35) and (38) together with $s_0^2 \geq 1$ are therefore necessary and sufficient conditions for stability and absence of vacuum Cherenkov radiation.

We now derive the condition for the group velocity of the linearized waves to be positive. This will play a role in the energy positivity constraint. In order for the waves to have positive group velocity, we must have $d\omega^2/dk^2 > 0$. In the limit $k \rightarrow \infty$, the group velocity is positive. Since $d\omega^2/dk^2$ is a continuous function of k , it is sufficient to show that there is no solution of the dispersion relation for which $d\omega^2/dk^2 = 0$.

Solving $d\omega^2/dk^2 = 0$ yields

$$2s_0^2k^2 = (\mu^2/c_{123} - m^2)s_0^2 + (1 + s_0^2)\omega^2. \quad (41)$$

Solving for k , and substituting back into the dispersion relation gives $a\omega^4 + b\omega^2 + c = 0$ where

$$a = (s_0^2 - 1)^2, \quad (42)$$

$$b = 2s_0^2[(1 - s_0^2)m^2 + (1 + s_0^2)\mu^2/c_{123} - 3\kappa_C\mu^2], \quad (43)$$

$$c = (m^2 - \mu^2/c_{123})^2 s_0^2. \quad (44)$$

Assuming (35) and $s_0^2 \geq 0$ implies $c \geq 0$. Just as in the discussion following (36), there are no real roots for ω (i.e. no positive roots for ω^2) if and only if $b > -2\sqrt{ac}$, where

$$b + 2\sqrt{ac} = 4s_0^2(1/c_{123} - 3\kappa_C/2)\mu^2. \quad (45)$$

Therefore the group velocity is positive for all k if and only if Eqs. (33) and (38) hold, as required by stability.

V. POSITIVITY OF ENERGY DENSITY

In addition to real frequencies, we further require that the linearized perturbations have positive energy. In doing so we necessarily run into the issue that there is no suitable local covariant expression for the energy density of a diffeomorphism-invariant theory that accounts for the energy in the gravitational field. However for wave-like linear perturbations that are periodic in time, it is possible to define the average energy density. The positive energy conditions for Einstein-aether theory have been found both by pseudotensor methods [21] and using the Noether current [16]. The two methods can be shown to produce equivalent results [22]; we will follow the latter approach, which is simpler both conceptually and computationally.

The Noether current is defined as follows [22]. We first define the canonical one-form θ^a via the total divergence that arises when varying the action

$$\delta S = \int \sqrt{|g|} d^4x (E[\psi] \cdot \delta\psi + \nabla_a \theta^a [\delta\psi]), \quad (46)$$

where E is the equation of motion, and ψ schematically denotes collectively all the dynamical fields. The Noether current 1-form $J^a[\xi]$ associated to the vector field ξ^a is then given by

$$J^a[\xi] = \theta^a[\mathcal{L}_\xi \psi] - L\xi^a, \quad (47)$$

where \mathcal{L}_ξ is the Lie derivative, and L is the Lagrangian. When the field equation $E = 0$ is satisfied the Noether current J^a constructed this way is conserved,

$$\nabla_a J^a = 0. \quad (48)$$

The energy is the Noether charge associated with the asymptotic time translation t^a . It is obtained by choosing ξ^a to coincide with t^a at infinity and integrating the flux of the corresponding current J^a through a Cauchy surface Σ ,

$$\mathcal{E} = \int_\Sigma J^a d\Sigma_a, \quad (49)$$

where $d\Sigma_a$ is the induced volume form on Σ .

The Noether current is a sum of two terms J_{∞} and J_φ which correspond to contributions from L_∞ and L_φ respectively. The Noether current J_∞ was found by Foster [23]. To evaluate J_φ it is useful to integrate the linearized Lagrangian by parts, giving a new Lagrangian

$$L'_\varphi = \frac{1}{2}\nabla_a\varphi\nabla^a\varphi - \frac{1}{2}m^2\varphi^2 + \mu Mu^a\nabla_a\varphi. \quad (50)$$

With suitable asymptotic boundary conditions this Lagrangian leads to the same equations of motion. It also gives the same Noether charge, and therefore the same time-averaged energy density, even though it does not give the same Noether current. The advantage of L'_φ over L_φ is that it contains no derivatives of u or g , so the corresponding canonical one-form contains only terms proportional to $\delta\varphi$, not to δu or δg . Varying the action, we find the contribution of L'_φ to the canonical one-form is

$$\theta_\varphi^a[\delta\varphi] = (\nabla^a\varphi + \mu Mu^a)\delta\varphi. \quad (51)$$

The corresponding Noether current is then determined from (47),

$$J_\varphi^a[\xi] = (\nabla^a\varphi + \mu Mu^a)\dot{\varphi} - L'_\varphi\xi^a. \quad (52)$$

We use the Noether current to find the energy density of linearized waves following Foster [16]. Consider a compact source in an asymptotically flat space-time. Fixing R to be a sphere of large coordinate radius r , the rate at which energy is radiated from the source is given by the flux of the Noether charge

$$-\dot{\mathcal{E}} = \int_R \vec{J} \cdot d\vec{A}, \quad (53)$$

where \vec{J} is the spatial part of J^a . The energy is being carried away by waves with average energy density u and group velocity v_g (which we assume positive), so that the average rate of energy loss is

$$-\langle \dot{\mathcal{E}} \rangle = \int_R u v_g dA, \quad (54)$$

where $\langle \cdot \rangle$ denotes time averaging over one period of the wave. Equating the two expressions for the rate of energy loss we find

$$u = \langle J_r \rangle / v_g. \quad (55)$$

To find the total energy density we need to evaluate this formula with the total Noether current, $J_\varphi + J_\infty$.

To carry out this calculation, note that for sufficiently large r all of the dynamical fields can be approximated by spherical waves for which

$$\partial_i\psi = -(1/v_p)\dot{\psi}\hat{r}_i, \quad (56)$$

where v_p is the phase velocity of the mode and \hat{r} is the outward facing unit normal. Equation (56) allows spatial derivatives in the energy density to be exchanged for

time derivatives. The contribution to the energy density coming from J_φ is

$$u_\varphi = \langle \dot{\varphi}^2 \rangle / (v_p v_g). \quad (57)$$

To find the total energy density we add this to the contribution to the energy density from the Lagrangian L_∞ [16]. Up to a positive factor we obtain

$$u \propto M^2(2 - c_{14})k^4 \langle \dot{f}^2 \rangle + 8c_{14} \langle \dot{\varphi}^2 \rangle. \quad (58)$$

It follows that $0 \leq c_{14} \leq 2$ is a sufficient condition for positive energy density. We cannot yet conclude that it is a necessary condition, since φ and f are related by the linearized equations of motion (24-27), which imply for the complex amplitudes

$$[c_{123}(\omega^2 - k^2 - m^2) + \mu^2] \varphi = \frac{1}{2}\mu M(1 - c_{13})i\omega k^2 f. \quad (59)$$

Substituting this into (58) we find that up to a positive multiplicative factor the energy density is given by

$$(2 - c_{14}) [c_{123}(\omega^2 - k^2 - m^2) + \mu^2]^2 + 2c_{14}\mu^2(1 - c_{13})^2\omega^2, \quad (60)$$

where ω and k are related by the dispersion relation (28). This expression holds for both of the spin-0 modes, each of which corresponds to a different solution of the dispersion relation.

We can now consider two separate limits of the dispersion relation. In the homogeneous limit $\omega = k = 0$, the second term in (60) drops out, and positivity requires $c_{14} \leq 2$. In the limit $k \rightarrow \infty$ with $\omega = k$, the first term is sub-leading in k , so that $c_{14} \geq 0$ is also needed. The upshot of this calculation is that the constraint from energy positivity of the spin-0 modes is

$$0 \leq c_{14} \leq 2, \quad (61)$$

which is the same as in Einstein-aether theory without the scalar coupling.

VI. COMBINED CONSTRAINTS

In this section we first gather together all the constraints derived from stability, Cherenkov radiation, and energy positivity. Next we consider the constraints from post-Newtonian effects, and finally combine these with the other constraints to determine the allowed parameter region.

A. Stability, Cherenkov and energy constraints

The constraints discussed already, together with those for the spin-1 and spin-2 modes [12], which are not af-

fected by the presence of the scalar field, are as follows:

$$\text{spin-0 stability } \frac{3\kappa_C}{2} \leq \frac{1}{c_{123}} \leq \frac{m^2}{\mu^2}, \quad (62)$$

$$\text{spin-0 speed } \frac{(2 - c_{14})c_{123}}{(1 - c_{13})(2 + c_{13} + 3c_2)c_{14}} \geq 1, \quad (63)$$

$$\text{spin-1 speed } \frac{2c_1 - c_1^2 + c_3^2}{2c_{14}(1 - c_{13})} \geq 1, \quad (64)$$

$$\text{spin-2 speed } \frac{1}{1 - c_{13}} \geq 1, \quad (65)$$

$$\text{spin-0 energy } 0 \leq c_{14} \leq 2, \quad (66)$$

$$\text{spin-1 energy } (2c_1 - c_1^2 + c_3^2)(1 - c_{13}) \geq 0, \quad (67)$$

$$\text{spin-2 energy } M^2 \geq 0. \quad (68)$$

Since the spin-1 and spin-2 modes satisfy a dispersion relation of the form $\omega^2 = s^2 k^2$, the stability constraint for each of these modes is implied by the Cherenkov constraint $s^2 \geq 1$. By contrast, the spin-0 modes have a k dependent phase velocity, so the separate stability and Cherenkov constraints involve several conditions.

The first inequality in (62) is actually implied by the other inequalities. To see this, note that (63), (65), and (66) imply that c_{123} and κ_C (6) must have the same sign. The first inequality in (62) therefore reduces to the inequality $c_{13} \leq 1$, which is already implied by (65).

B. Post-Newtonian parameters

As we mentioned in Sec. I, Einstein-aether theory has a Newtonian limit with a renormalized gravitational constant G_N given in (5). The spin-2 energy constraint says that G must be positive. Combined with the spin-0 energy constraint, this also implies that κ_N and therefore Newton's constant G_N (5) is also positive.

The post-Newtonian parameters β and γ take the same values (unity) as in general relativity [24]. The effects of Lorentz violation are captured at first post-Newtonian order by the dimensionless preferred frame parameters α_1 and α_2 . These can be expressed in terms of $c_{1,2,3,4}$ as [25, 26]

$$\alpha_1 = \frac{-8(c_3^2 + c_1 c_4)}{2c_1 - c_1^2 + c_3^2}, \quad (69)$$

$$\alpha_2 = \frac{\alpha_1 - (c_1 + 2c_3 - c_4)(2c_1 + 3c_2 + c_{34})}{2c_{123}(2 - c_{14})}. \quad (70)$$

The preferred frame parameters are constrained observationally to be small [27], $\alpha_1 \lesssim 10^{-4}$ and $\alpha_2 \lesssim 4 \times 10^{-7}$, and because Einstein-aether theory has four free parameters they can be made to vanish exactly in a two-dimensional subspace of parameter space. We will see that similar considerations apply also in the presence of the scalar field.

We now consider the effect on the post-Newtonian expansion of coupling to the scalar field φ . The stability

constraint (62) shows that φ cannot be massless if the cosmological Newton constant is to be positive. In order to simplify our analysis, we will further assume that m is large enough so that the Compton wavelength of φ particles is much smaller than any scale on which gravity has been tested in the late universe. Newton's law of gravitation has been tested to sub-millimeter scales, so we shall assume the lower bound $m \gtrsim 10^{-3}$ eV on the mass. If the mass is smaller than this, a more complete analysis would be required to determine the effects of the scalar field coupling to gravity. Most likely the mass of the inflaton must be tremendously larger than this anyway.

Assuming that all fields vary on scales much larger than $1/m$, we can integrate out the scalar field. We first solve the linearized scalar field equation (9) by an expansion in \square/m^2 ,

$$\varphi = -(\square + m^2)^{-1} \mu M \theta = -\frac{\mu M}{m^2} \theta + \frac{\mu M}{m^4} \square \theta + \dots, \quad (71)$$

where in what follows we will keep only the leading order contribution, $\varphi = -(\mu M/m^2)\theta$. If we substitute this value back into the linearized scalar field action and expand to $O(\varphi^2)$ we find

$$L_\varphi = \frac{\mu^2 M^2}{2m^4} \nabla_a \theta \nabla^a \theta + \frac{\mu^2 M^2}{2m^2} \theta^2. \quad (72)$$

The $(\nabla \theta)^2$ term is of higher order in the derivative expansion, and so will be neglected. The remaining term can be absorbed in the c_2 term of the aether action, which is given by $-\frac{1}{2}M^2 c_2 \theta^2$. Therefore, upon integrating out the scalar field, the new Lagrangian is simply L_a with a new value of the parameter c_2 which we denote c'_2 ,

$$c_2 \rightarrow c'_2 = c_2 - \frac{\mu^2}{m^2}. \quad (73)$$

The higher derivative terms in the expansion (71) lead to higher derivative terms in the aether action. If we neglect these higher-order corrections, the post-Newtonian parameters can be determined from the known results for Einstein-aether theory (69) and (70) using c'_2 in place of c_2 .

Just as in pure Einstein-aether theory, we can set both preferred frame parameters α_1, α_2 to zero, now by choosing

$$c_2 = \frac{-2c_1^2 - c_1 c_3 + c_3^2}{3c_1} + \frac{\mu^2}{m^2}, \quad (74)$$

$$c_4 = -\frac{c_3^2}{c_1}. \quad (75)$$

Once this choice of c_4 has been made, the spin-1 and spin-2 speed constraints and the spin-0 and spin-1 energy constraints (64)-(67) can be conveniently expressed in terms of $c_\pm \equiv c_1 \pm c_3$, and they hold if and only if $0 \leq c_+ \leq 1$ and $c_- \geq 0$ [26]. Moreover, once this choice of c_2 has been made, we have

$$c_{123} = \frac{c_{13}^2}{3c_1} + \frac{\mu^2}{m^2}, \quad (76)$$

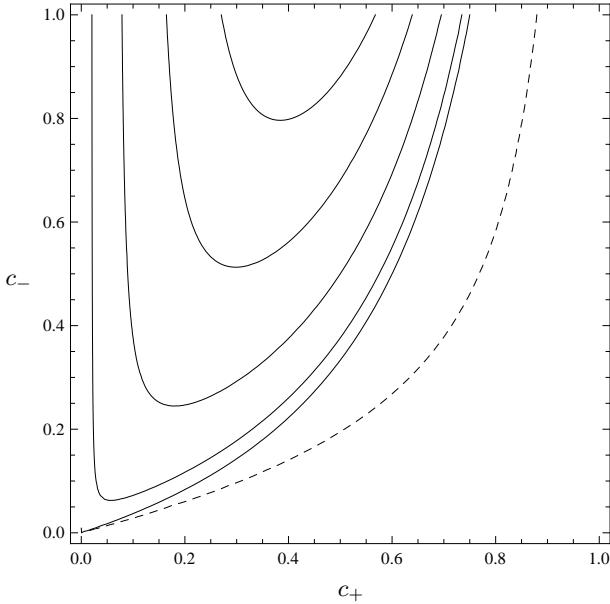


FIG. 1: The allowed ranges for the parameters $c_{\pm} = c_1 \pm c_3$, assuming c_2 and c_4 are chosen so that the preferred frame PPN parameters vanish. The spin-0 Cherenkov constraint is satisfied below the solid lines, which correspond (from lowest to highest) to $\mu = 0$, $\mu = \frac{1}{4}m$, $\mu = \frac{1}{2}m$, $\mu = m$, and $\mu/m \rightarrow \infty$. The dashed line corresponds to the parameter values for which the damping rate from weak-field binary systems calculated in [16] agrees with general relativity.

which is positive when the other constraints hold, since the positivity of c_1 is implied by $c_{\pm} \geq 0$. The second inequality of (62) then becomes

$$\mu^2 \leq c_{123}m^2 = \mu^2 + \frac{c_{13}^2}{3c_1}m^2, \quad (77)$$

which is also automatically satisfied when the other constraints are imposed.

The condition for absence of vacuum Cherenkov radiation into the spin-0 mode (63) depends on c_2 , and therefore on μ via (74). In pure Einstein-aether theory (i.e. when $\mu = 0$), the value of c_2 in Eq. (74) goes to zero as $c_{\pm} \rightarrow 0$. In this limit, s_0 (29) depends on the ratio between c_+ and c_- . However, if $\mu \neq 0$, then c_2 does not go to zero, and $s_0 \rightarrow \infty$ as $c_{\pm} \rightarrow 0$. This means that the Cherenkov constraint can always be satisfied in an open neighborhood of $c_{\pm} = 0$. As shown in Fig. 1, the range of allowed c_{\pm} is significantly enlarged as μ increases.

C. Other constraints

Further constraints arise from the effective cosmological value of Newton's constant in the Friedmann equation, and radiation and orbital dynamics of compact binaries.

As explained above, when the PPN parameters $\alpha_{1,2}$ have been set to zero by the choices (74) and (75), it follows from the other constraints that $c_{123} \geq 0$. Then (63) also implies that κ_C (6) is positive. This would be the value of the cosmological Newton constant relevant during inflation. For later times, after the inflaton is integrated out¹, the relevant value is $G'_C = \kappa'_C G$, where κ'_C is given by (6) but with c'_2 in place of c_2 . When c_2 and c_4 are set according to (74,75) it follows [26] that $\kappa'_C = \kappa_N$. Predictions of primordial element abundances from big bang nucleosynthesis are sensitive to G'_C , and agreement with observations requires [14] that G'_C not differ too much from earthbound measurements of Newton's constant,

$$|G'_C/G_N - 1| \lesssim 1/8. \quad (78)$$

This constraint is thus trivially satisfied when the preferred frame PPN parameters vanish.

The rate of change in orbital period of binary pulsars due to gravitational and aether radiation, ignoring effects from strong self-gravity of the bodies, was computed by Foster [16]. The observed orbital decay agrees with that found in general relativity to within better than one percent [27]. The constraint that the rate in weak field Einstein-aether theory agree with that found in general relativity is indicated in Fig. 1. The strong-field effects that contribute to the damping rate and orbital dynamics of binary systems was computed in [28] in terms of as yet unknown parameters that characterize the velocity dependence of the energy (or action) of the bodies. Strong-field effects are likely to lead to further constraints on the parameters, restricting c_{\pm} to be less than $\sim 0.01 - 0.1$.

VII. HOMOGENEOUS COSMOLOGY

We consider now the cosmology of Einstein-aether theory coupled to a single scalar inflaton field, with a potential that depends on the expansion of the aether. The metric is assumed to be homogeneous, isotropic, and spatially flat, hence it can be written in the form

$$ds^2 = dt^2 - a(t)^2 d\vec{x}^2. \quad (79)$$

In a homogeneous and isotropic solution the aether must be aligned with the cosmological rest frame. In this case, the spatial components of the aether equation are automatically satisfied, and the time component of the aether equation of motion (10) just determines the Lagrange multiplier λ . That the aether relaxes to such an isotropic configuration during cosmological expansion is

¹ The derivative terms in the Klein-Gordon equation are presumably negligible when $H \ll m$, which occurs at a temperature $T \ll \sqrt{mM}$. This is satisfied at the nucleosynthesis temperature ~ 0.1 MeV as long as $m \gg 10^{-18}$ eV.

shown perturbatively in [3, 7]. A nonlinear extension of this analysis is currently underway [29].

We note that our results for homogeneous cosmology apply as well to the extended version of Hořava gravity [30] proposed in Ref. [31]. As shown in Ref. [32], this follows because the aether in these solutions is hypersurface-orthogonal.

With the metric (79), the expansion is $\theta = 3H$ and the homogeneous scalar (inflaton) field equation (9) is

$$\ddot{\varphi} + \theta\dot{\varphi} + V_\varphi = 0, \quad (80)$$

where the driving force $-V_\varphi$ is now dependent on the expansion θ . The terms in V_φ containing positive powers of θ lead to driving forces that are most relevant during inflation when the expansion is large. As the expansion slows, these forces subside, allowing for a graceful end to inflation.

The Friedmann equation can be derived by considering a metric of the form $ds^2 = N(t)^2 dt^2 - a(t)^2 d\vec{x}^2$ and varying the action with respect to N . The normalization condition for u together with the symmetry completely fixes u to $u^a = (1/N, 0, 0, 0)$, and the symmetry reduced Lagrangian density is (up to a total derivative)

$$\sqrt{-g}L = Na^3 \left(-\frac{M_C^2}{3}\theta^2 + \frac{1}{2N^2}\dot{\varphi}^2 - V(\theta, \varphi) \right), \quad (81)$$

where $\theta = 3H/N$ and we have introduced a “reduced cosmological Planck mass” $M_C \equiv M/\sqrt{\kappa_C}$ with κ_C defined in (6). For terms in the action that are linear in θ , the factor of $1/N$ in θ cancels the factor N in the determinant of the metric, so that such terms do not contribute to the Friedmann equation. This is in agreement with the general form of the stress tensor (14).

Varying (81) with respect to N gives the Friedmann equation

$$\theta^2 = \frac{3}{M_C^2} \left(\frac{1}{2}\dot{\varphi}^2 + V - \theta V_\theta \right), \quad (82)$$

where the gauge condition $N = 1$ has been adopted. This equation determines θ only implicitly as a function of φ and $\dot{\varphi}$, since the potential V also depends on θ .

We now focus on the case when the potential has only quadratic terms,

$$V(\theta, \varphi) = \frac{1}{2}m^2\varphi^2 + \mu M\theta\varphi, \quad (83)$$

with $\mu > 0$. The Friedmann equation (82) then has the standard form for a massive scalar. While the $\theta\varphi$ coupling does not show up in the Friedmann equation, it does affect the Einstein equation, via the pressure term in the scalar stress tensor (14),

$$p = \dot{V}_\theta = \mu M\dot{\varphi}. \quad (84)$$

The inflaton field equation (80) becomes

$$\ddot{\varphi} + \theta\dot{\varphi} + m^2\varphi + \mu M\theta = 0. \quad (85)$$

The last term acts as an external force that pushes φ in the negative direction.

Slow roll solutions can be obtained by neglecting the kinetic term $\dot{\varphi}^2$ in (82), and the term $\ddot{\varphi}$ in the Klein-Gordon equation (80). The slow roll equations in this case become

$$\theta = \sqrt{\frac{3}{2}} \frac{m}{M_C} |\varphi|, \quad (86)$$

$$\dot{\varphi} = -\sqrt{\frac{2}{3}} m M_C \left(\text{sgn}(\varphi) + \frac{\mu}{\mu_c} \right), \quad (87)$$

where we have defined

$$\mu_c = \sqrt{\frac{2}{3\kappa_C}} m = \sqrt{\frac{2 + c_{13} + 3c_2}{3}} m. \quad (88)$$

This slow roll solution need not be stable: If the Hubble force $-\mu M\theta$ is dominant then there is a feedback effect, where this force leads to larger field values, which leads to a more rapid expansion and therefore a stronger Hubble force. The slow roll solution will be stable provided that φ and $\dot{\varphi}$ are of the opposite sign for all values of φ , i.e. if $|\mu| \leq \mu_c$. This condition is implied by the condition for stability of the linearized modes (62). This is not surprising, since that condition was originally inferred from the zero wavevector limit of the dispersion relation (34), and we have restricted even in the nonlinear analysis to a quadratic potential.

The general behavior of this dynamical system for different values of μ/μ_c is illustrated with phase portraits in Fig. 2. Plotted there is the flow on the $(\varphi, \dot{\varphi})$ plane that is obtained when θ is eliminated using the Friedmann equation. In terms of the dimensionless variables $\tilde{\varphi} = \varphi/M_C$ and $\tilde{t} = mt$, we can write the system in the form

$$\frac{d}{d\tilde{t}}(\tilde{\varphi}, \tilde{\varphi}') = \left(\tilde{\varphi}', -\tilde{\varphi} - \sqrt{\frac{3}{2}(\tilde{\varphi}'^2 + \tilde{\varphi}^2)} \left(\tilde{\varphi}' + \sqrt{\frac{2}{3}} \frac{\mu}{\mu_c} \right) \right), \quad (89)$$

where the prime denotes derivative with respect to \tilde{t} . The right hand side is the vector field plotted in the figure.

The number of e-folds of expansion during inflation is $N = \ln(a_f/a_i) = \int da/a$, where a_i and a_f are the scale factors at the beginning and end of inflation, respectively. N can be computed in terms of the field values at the start and end of inflation as

$$N = \frac{1}{3} \int \theta dt = \frac{1}{3} \int \frac{\theta}{\dot{\varphi}} d\varphi. \quad (90)$$

During slow-roll θ and $\dot{\varphi}$ are given by (86) and (87), and φ has a fixed sign, so

$$N = \frac{\varphi_i^2 - \varphi_f^2}{4M_C^2(1 + \text{sgn}(\varphi_i)\mu/\mu_c)}. \quad (91)$$

where φ_i and φ_f are the values of the inflaton field at the start and end of inflation, respectively. As $\mu \rightarrow \mu_c$

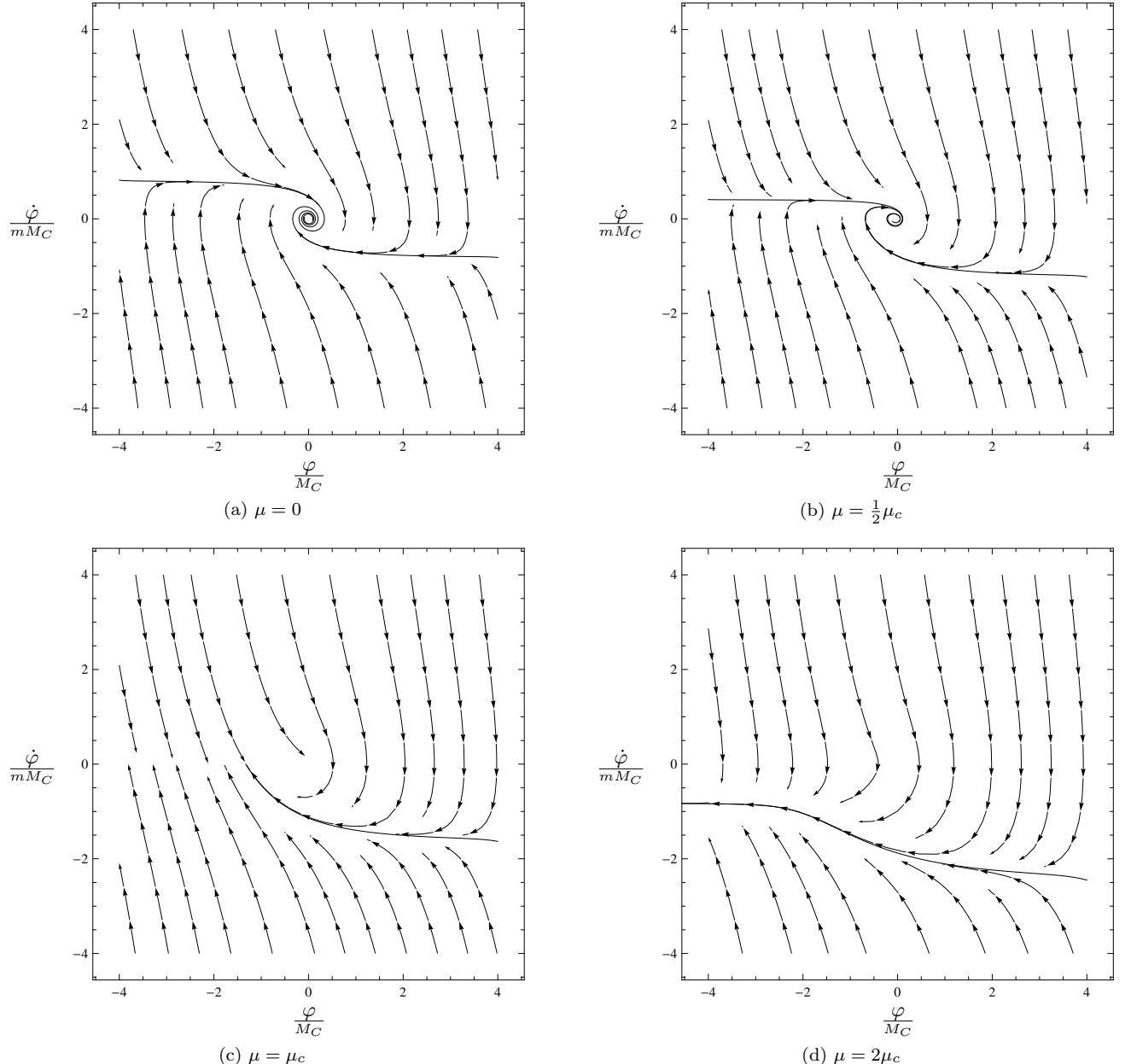


FIG. 2: Phase portraits of inflaton dynamics with the potential $V = \frac{1}{2}m^2\varphi^2 + \mu M\theta\varphi$, given by the dynamical system (89) with $\mu > 0$. (2a) The result for standard slow-roll inflation. Solutions are attracted toward the slow roll solution, which appears as two almost-horizontal lines on the plot. At the end of the slow roll period, a reheating phase begins where the field undergoes damped oscillation about the minimum of the potential, which appears as an inward spiral in the phase portrait. In (2b) the basic picture remains the same, but the value of $\dot{\varphi}$ for the slow roll solutions is changed. The duration of the slow roll period is increased if $\varphi_i < 0$ or decreased if $\varphi_i > 0$. In (2c), μ is at exactly μ_c , so any configuration with $\varphi < 0$ and $\dot{\varphi} = 0$ is a fixed point, corresponding to a de Sitter space-time.

When μ is increased beyond the critical value, as in (2d), the slow roll solution becomes unstable and φ grows without bound. Although only portraits with $\mu > 0$ have been shown, the system has combined $\mu \rightarrow -\mu$ and $\varphi \rightarrow -\varphi$ symmetry, allowing the phase portraits with $\mu < 0$ to be obtained from the phase portraits with $\mu > 0$ by a reflection through the origin.

one of the slow roll solutions approaches a de Sitter solution with constant negative φ , corresponding to the limit $N \rightarrow \infty$. These solutions are illustrated in Fig. 2c on the negative φ axis.

The number of e-folds can be made arbitrarily large compared to the $\mu = 0$ case, for fixed values of $\varphi_{i,f}$ and M_C , by taking $\mu \rightarrow \mu_c$. However, the ratio μ/μ_c is constrained by the second inequality in (62), which is the condition that the spin-0 aether mode not grow exponentially in the long-wavelength limit. Expressed in terms of μ_c (88), it yields

$$\frac{\mu^2}{\mu_c^2} \leq \frac{3c_{13} + 3c_2}{2 + c_{13} + 3c_2}. \quad (92)$$

It follows that the limit $\mu \rightarrow \mu_c$ is achieved only by $c_{13} \rightarrow 1$ or by $c_2 \rightarrow \infty$. While the possibility that c_{13} could be close to one has not been ruled out, all the mode speeds (63)-(65) diverge in this limit, suggesting that it is pathological. Additionally, it is expected that a strong-field analysis of binary systems will constrain c_{\pm} to be smaller than order unity. As for the case where $c_2 \rightarrow \infty$, as noted in Sec. VI B, φ can be integrated out so that the low-energy theory depends only on c'_2 , given by (73). It is possible to keep c'_2 small even while c_2 becomes large by taking μ/m large, though this requires a cancellation to occur between c_2 and μ^2/m^2 .

If the preferred-frame parameters α_1 and α_2 are constrained to vanish exactly as in Sec. VI B, then as explained above, the constraint (92) is implied by the other constraints. In this case, c_2 and therefore μ_c are determined by μ/m and c_{\pm} . Then μ/μ_c can be expressed as

$$\frac{\mu^2}{\mu_c^2} = \left(1 + \frac{2m^2 c_+ + c_- (1 - c_+)}{3\mu^2 c_+ + c_-} \right)^{-1}. \quad (93)$$

This approaches 1 only if $\mu/m \rightarrow \infty$, or $c_+ \rightarrow 1$ and $c_- \rightarrow \infty$. These are the same conditions as inferred above without having constrained $\alpha_{1,2}$ to vanish.

A. Cosmological perturbations

Finally, we briefly discuss the effect of inflaton-aether coupling on the spectrum of perturbations generated during inflation. In single-field slow roll inflation the primordial perturbations originate from quantum vacuum fluctuations of the coupled inflaton-metric mode. The introduction of the aether field leads to an additional spin-0 as well as a spin-1 mode, each with their own quantum fluctuations. If these modes are not directly coupled to matter and the background is exactly de Sitter, the aether modes decay exponentially [3], though they may be sourced during reheating by anisotropic stresses [4]. These perturbations may grow in power-law space-times for certain values of the parameters $c_{1,2,3,4}$ [6].

Introduction of scalar-aether coupling further modifies this scenario. There are two spin-0 coupled inflaton-aether-metric modes whose dispersion relation is nontriv-

ial even in flat space. The generalization of this dispersion relation to curved space-time determines both the amplitude of vacuum fluctuations and the time at which the modes freeze. Moreover, the modified background equation of motion determines how the Hubble parameter changes with time, and therefore the time at which modes of different comoving wavelengths reach the Hubble radius.

A priori it is not clear whether the spin-0 modes will freeze or decay on superhorizon scales. This behavior is determined by how the modes couple to the background expansion, and is different for a scalar and the spin-0 part of a vector.

VIII. SUMMARY

We have considered Einstein-aether theory coupled to a scalar field via a potential that depends on the local rate of expansion of space in the frame of the aether, $\theta = \nabla_a u^a$. This could be an effective field theory description of Lorentz violating UV physics in the vacuum, for example, at the scale of a fundamental cutoff. We have mostly focused on the lowest order term $\mu M \theta \varphi$, which leads to breaking of time-reversal invariance, non-trivial dispersion, and a cosmological dynamics modified by a driving force on the scalar proportional to θ .

Although we are most interested in this model for the potentially observable effects of Lorentz violation on cosmology, we have strived to examine all of the theoretical and observational constraints that arise when perturbing around a locally flat space-time. We find that when the post-Newtonian parameters are matched to general relativity, the combined constraints on the remaining free couplings c_1 and c_3 imposed by stability, positive energy density and absence of vacuum Cherenkov radiation, take the same form as in Einstein-aether theory, except for the spin-0 Cherenkov constraint, which is relaxed (Fig. 1). There is a single constraint on the new parameter μ , the second inequality of (62), which is automatically satisfied when the PPN parameters match those of general relativity (77).

An aether field uncoupled to matter does not affect the dynamics of homogeneous, isotropic, spatially flat cosmology except by a renormalization of Newton's constant $G \rightarrow G_C$ (6). It turns out that the lowest order scalar-aether coupling $\theta \varphi$ contributes isotropic pressure and no energy density. Therefore it does not affect the Friedmann equation, but the driving force it adds to the scalar field equation can act either to oppose or accelerate the expansion of the universe in a slow-roll scenario, altering the number of e-folds of expansion. An unbounded increase in the number of e-folds is possible, although this requires carefully chosen values of the couplings in order to evade current constraints, and may be ruled out by future constraints. This illustrates the importance of considering all constraints when building cosmological models.

A next step would be to derive the spectrum of primordial perturbations in the model. This would involve generalizing the treatment of linearized perturbations from flat space to a general inflating background. This differs from previous work since the two scalar modes—that of the inflaton and that of the aether—are mixed, and because the slow-roll dynamics is modified. In particular, it would be interesting to determine how the spectral index and the tensor-to-scalar ratio depend on the coupling constants of the theory.

While we have focused on the coupling to a scalar inflaton, other types of coupling to the expansion are possible, allowing for a wide range of Lorentz-violating phenomena. For example, any scalar operator could appear in the Lagrangian multiplied by θ , allowing effects to be switched off (or on) as the expansion of the universe

slows. In this context, it should be kept in mind that in the present universe the largest potentially observable values of θ are not due to the Hubble expansion, but rather probably occur outside stellar mass black holes, where $\theta \sim \text{km}^{-1}$.

Acknowledgments

The authors thank Brendan Foster and Eugene Lim for helpful discussions. This research was supported in part by the Foundational Questions Institute (FQXi grant RFP20816), NSF Grants No. 0601800 and No. 0903572, and by an NSERC PGS-D to WD.

-
- [1] D. Mattingly, “Modern Tests of Lorentz Invariance,” *Living Reviews in Relativity* **8** no. 5, (2005) .
<http://www.livingreviews.org/lrr-2005-5>.
 - [2] M. Gasperini, “Inflation and broken lorentz symmetry in the very early universe,” *Physics Letters B* **163** no. 1-4, (1985) 84 – 86.
 - [3] E. A. Lim, “Can we see Lorentz-violating vector fields in the CMB?,” *Phys. Rev.* **D71** (2005) 063504, [arXiv:astro-ph/0407437](https://arxiv.org/abs/astro-ph/0407437).
 - [4] B. Li, D. Fonseca Mota, and J. D. Barrow, “Detecting a Lorentz-Violating Field in Cosmology,” *Phys. Rev.* **D77** (2008) 024032, [arXiv:0709.4581 \[astro-ph\]](https://arxiv.org/abs/0709.4581).
 - [5] J. A. Zuntz, P. G. Ferreira, and T. G. Zlosnik, “Constraining Lorentz violation with cosmology,” *Phys. Rev. Lett.* **101** (2008) 261102, [arXiv:0808.1824 \[gr-qc\]](https://arxiv.org/abs/0808.1824).
 - [6] C. Armendariz-Picon, N. F. Sierra, and J. Garriga, “Primordial Perturbations in Einstein-Aether and BPSH Theories,” *JCAP* **1007** (2010) 010, [arXiv:1003.1283 \[astro-ph.CO\]](https://arxiv.org/abs/1003.1283).
 - [7] S. Kanno and J. Soda, “Lorentz violating inflation,” *Phys. Rev.* **D74** (2006) 063505, [arXiv:hep-th/0604192](https://arxiv.org/abs/hep-th/0604192).
 - [8] P. Avelino, D. Bazeia, L. Losano, R. Menezes, and J. Rodrigues, “Impact of Lorentz violation on the dynamics of inflation,” *Phys. Rev.* **D79** (2009) 123503, [arXiv:0903.5297 \[astro-ph.CO\]](https://arxiv.org/abs/0903.5297).
 - [9] F. R. Klinkhamer and G. E. Volovik, “Towards a solution of the cosmological constant problem,” *JETP Lett.* **91** (2010) 259–265, [arXiv:0907.4887 \[hep-th\]](https://arxiv.org/abs/0907.4887).
 - [10] T. G. Zlosnik, P. G. Ferreira, and G. D. Starkman, “Modifying gravity with the Aether: an alternative to Dark Matter,” *Phys. Rev.* **D75** (2007) 044017, [arXiv:astro-ph/0607411](https://arxiv.org/abs/astro-ph/0607411).
 - [11] J. Zuntz, T. G. Zlosnik, F. Bourliot, P. G. Ferreira, and G. D. Starkman, “Vector field models of modified gravity and the dark sector,” *Phys. Rev.* **D81** (2010) 104015, [arXiv:1002.0849 \[astro-ph.CO\]](https://arxiv.org/abs/1002.0849).
 - [12] T. Jacobson, “Einstein-aether gravity: a status report,” *Pos QG-PH* (2007) 020, [arXiv:0801.1547 \[gr-qc\]](https://arxiv.org/abs/0801.1547).
 - [13] D. Mattingly and T. Jacobson, “Relativistic gravity with a dynamical preferred frame,” in *Proceedings of the Second Meeting on CPT and Lorentz Symmetry*, V. A. Kostelecky, ed. August, 2001. [arXiv:gr-qc/0112012](https://arxiv.org/abs/gr-qc/0112012).
 - [14] S. M. Carroll and E. A. Lim, “Lorentz-violating vector fields slow the universe down,” *Phys. Rev.* **D70** (2004) 123525, [arXiv:hep-th/0407149](https://arxiv.org/abs/hep-th/0407149).
 - [15] C. Eling and T. Jacobson, “Spherical Solutions in Einstein-Aether Theory: Static Aether and Stars,” *Class. Quant. Grav.* **23** (2006) 5625–5642, [arXiv:gr-qc/0603058](https://arxiv.org/abs/gr-qc/0603058).
 - [16] B. Z. Foster, “Radiation Damping in Einstein-Aether Theory,” *Phys. Rev.* **D73** (2006) 104012, [arXiv:gr-qc/0602004](https://arxiv.org/abs/gr-qc/0602004).
 - [17] T. Jacobson and D. Mattingly, “Einstein-aether waves,” *Phys. Rev.* **D70** (2004) 024003, [arXiv:gr-qc/0402005](https://arxiv.org/abs/gr-qc/0402005).
 - [18] S. M. Carroll, T. R. Dulaney, M. I. Gresham, and H. Tam, “Instabilities in the Aether,” *Phys. Rev.* **D79** (2009) 065011, [arXiv:0812.1049 \[hep-th\]](https://arxiv.org/abs/0812.1049).
 - [19] W. Donnelly and T. Jacobson, “Stability of the aether,” *Phys. Rev.* **D82** (2010) 081501, [arXiv:1008.4351 \[hep-th\]](https://arxiv.org/abs/1008.4351).
 - [20] J. W. Elliott, G. D. Moore, and H. Stoica, “Constraining the new aether: Gravitational Cherenkov radiation,” *JHEP* **08** (2005) 066, [arXiv:hep-ph/0505211](https://arxiv.org/abs/hep-ph/0505211).
 - [21] C. Eling, “Energy in the Einstein-aether theory,” *Phys. Rev.* **D73** (2006) 084026, [arXiv:gr-qc/0507059](https://arxiv.org/abs/gr-qc/0507059).
 - [22] V. Iyer and R. M. Wald, “Some properties of Noether charge and a proposal for dynamical black hole entropy,” *Phys. Rev.* **D50** (1994) 846–864, [arXiv:gr-qc/9403028](https://arxiv.org/abs/gr-qc/9403028).
 - [23] B. Z. Foster, “Noether charges and black hole mechanics in Einstein- aether theory,” *Phys. Rev.* **D73** (2006) 024005, [arXiv:gr-qc/0509121](https://arxiv.org/abs/gr-qc/0509121).
 - [24] C. Eling and T. Jacobson, “Static post-Newtonian equivalence of GR and gravity with a dynamical preferred frame,” *Phys. Rev.* **D69** (2004) 064005, [arXiv:gr-qc/0310044](https://arxiv.org/abs/gr-qc/0310044).
 - [25] M. L. Graesser, A. Jenkins, and M. B. Wise, “Spontaneous Lorentz violation and the long-range gravitational preferred-frame effect,” *Phys. Lett.* **B613** (2005) 5–10, [arXiv:hep-th/0501223](https://arxiv.org/abs/hep-th/0501223).
 - [26] B. Z. Foster and T. Jacobson, “Post-Newtonian parameters and constraints on Einstein- aether theory,”

- Phys. Rev.* **D73** (2006) 064015, [arXiv:gr-qc/0509083](#).
- [27] C. M. Will, “The confrontation between general relativity and experiment,” *Living Rev. Rel.* **9** (2005) 3, [arXiv:gr-qc/0510072](#).
- [28] B. Z. Foster, “Strong field effects on binary systems in Einstein-aether theory,” *Phys. Rev.* **D76** (2007) 084033, [arXiv:0706.0704 \[gr-qc\]](#).
- [29] I. Carruthers and T. Jacobson, “Cosmic alignment of the aether,” *Phys. Rev.* **D83** (2011) 024034, [arXiv:1011.6466 \[gr-qc\]](#).
- [30] P. Horava, “Quantum Gravity at a Lifshitz Point,” *Phys. Rev.* **D79** (2009) 084008, [arXiv:0901.3775 \[hep-th\]](#).
- [31] D. Blas, O. Pujolas, and S. Sibiryakov, “Consistent Extension of Horava Gravity,” *Phys. Rev. Lett.* **104** (2010) 181302, [arXiv:0909.3525 \[hep-th\]](#).
- [32] T. Jacobson, “Extended Horava gravity and Einstein-aether theory,” *Phys. Rev.* **D81** (2010) 101502, [arXiv:1001.4823 \[hep-th\]](#).